

### Egozentrierte Netzwerke in der Forschungspraxis: Ausschöpfungsquoten und Validität soziodemographischer Variablen

Schenk, Michael; Mohler, Peter Ph.; Pfennig, Uwe

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Schenk, M., Mohler, P. P., & Pfennig, U. (1992). Egozentrierte Netzwerke in der Forschungspraxis: Ausschöpfungsquoten und Validität soziodemographischer Variablen. *ZUMA Nachrichten*, 16(31), 87-120. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-209668>

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# EGOZENTRIERTE NETZWERKE IN DER FORSCHUNGSPRAXIS:

## Ausschöpfungsquoten und Validität soziodemographischer Variablen

*Michael Schenk, Peter Ph. Mohler und Uwe Pfenning, unter  
Mitwirkung von Renate Ell*

Der folgende Beitrag knüpft an die bisherigen Beiträge in den ZUMA-Nachrichten zur sozialwissenschaftlichen Netzwerkforschung an. Diese haben sich, basierend auf einer von Prof. Dr. Hartmut Esser bei ZUMA initiierten, lokalen Netzwerkstudie, mit Fragen des Vergleichs von Namensgeneratoren und der Zuverlässigkeit von Netzwerkangaben aus der Perspektive der befragten Ego-Person beschäftigt. Ausstehend sind Analysen über die Validität von Ego-Informantenangaben. Die nachfolgenden Analysen aus einer DFG-Studie an der Universität Hohenheim, die von Prof. Dr. Michael Schenk beantragt wurde, haben diesen Validierungsaspekt zum Inhalt. Die Besonderheit dieser Studie liegt darin, daß erstmals auch in (relativ) größerem Umfang die von einer Zielperson (EGO) genannten Netzpersonen (ALTERI) in sogenannten Follow-Up-Interviews befragt wurden. Darüber hinaus wurden auch die Zielpersonen zweimalig befragt, so daß einfache Test-Retest-Analysen zu Eigenangaben im Vergleich zur Validität der Angaben über Netzpersonen möglich sind. Der Validitätsbegriff bezieht sich hierbei auf die Übereinstimmung der Angaben von Ego über "seine" Netzpersonen (Proxy-Daten) mit den Angaben der Netzpersonen selbst. Zu Beginn des Beitrages werden zudem stichprobentheoretische Aspekte von Netzwerkstudien mit Follow-Up-Interviews dargestellt. Dieser Teil der Analyse hat die Untersuchung der Praktikabilität und möglicher Verzerrungseffekte zum Inhalt.

## 1. Egozentrierte Netzwerke in der Forschungspraxis

Die Erhebung egozentrierter Netzwerke basiert auf der Verfügbarkeit von Informationen über die von einer Zielperson genannten Netzpersonen. Der Umfang dieser Informationen reicht von Angaben über die einzelnen Beziehungen der Zielperson zu jeder Netzperson (Ego-Alter-Dyade) bis zu Angaben über die einzelnen Beziehungen der Netzpersonen untereinander (Alter-Alter-Relationen). Aus den Angaben der Ego-Alter-Dyaden können bereits einzelne Determinanten zur Analyse sozialer egozentrierter Netzwerke berechnet werden. Unter anderem sind dies Heterogenitäts- und Homogenitätsmaße sowie spezifische Kommunikationsrollen wie beispielsweise Meinungsführer und Empfänger. Die Angaben über die Alter-Alter-Relationen werden benötigt zur Bestimmung der Netzwerkgeschlossenheit und der Netzwerkdichte, also zu zentralen beschreibenden Parametern der Netzstruktur. Geschlossenheit bezeichnet hierbei das Verhältnis der tatsächlichen Alteri-Kontakte zu den potentiell möglichen Alteri-Kontakten. Der Begriff der Netzwerkdichte dient der Beschreibung einer generalisierten sozialen Harmonie unter den verschiedenen Alteri. Daraus lassen sich Analysen zur Mikrostruktur des egozentrierten Netzwerkes, zu den die Netzwerkanalyse charakterisierenden Auswertungen über Triaden und Trippletts, zu Cliques und Brückenfunktionen sowie zur Balance-Theorie sozialer Beziehungen gewinnen.

Letztlich kann nur bei Erhebung beider Beziehungsstrukturen, also der Ego-Alter-Dyaden und der Alter-Alter-Dyaden von einem "echten" egozentrierten Netzwerk gesprochen werden. Denn gemäß der oft zitierten Definition von sozialen Netzwerken im Sinne von John C. Mitchell (1969:2), die die Einbeziehung der Sozialstruktur in die Analyse hervorhebt - "(...) a specific set of linkages between a defined set of persons, with the additional property that the characteristics of the linkages as a whole may be used to interpret the social behavior of the persons involved" - werden die Angaben über die globalen Netzparameter benötigt. Die Verfügbarkeit beider Informationen erlaubt also erst eine definitorisch genügende Analyse von sozialen Netzwerken.

Der Zugang zu diesen beiden Informationsquellen ist mit erheblichen methodischen Problemen verbunden. Zum einen können die Informationen aus einer Befragung der genannten Netzpersonen gewonnen werden. Dieses, bereits von Elihu Katz und Paul F. Lazarsfeld in der bekannten Decatur-Studie über "Personal Influence" (Katz/Lazarsfeld 1955, 1965) angewandte Verfahren der sogenannten Follow-Up-Interviews, bedingt die Verfügbarkeit von Adressen der genannten Netzpersonen. Zum anderen kann man/frau die Angaben allein von Ego erheben und die Zielperson sozusagen als Informant benutzen. In

der methodischen Literatur wird dieses Verfahren als Proxy-Interview bezeichnet, da es sich um Schätzungen, um erwartete und vermutete Merkmale von Drittpersonen handelt.

Damit werden die Probleme der beiden Erhebungsverfahren offenbar. Follow-Up-Interviews führen zu einem hohen Kostenaufwand und bedürfen eines aufwendigen Adressenmanagements. In Zeiten hoher Datenschutzanforderungen und hoher Interviewkosten stößt dieses Verfahren schnell an (s)eine praktische Grenze. Hinzu kommt die Sensibilisierung der befragten Personen hinsichtlich der Bereitschaft zur Adressennennung. Dies hat methodische Folgen: Es gilt zu überprüfen, inwieweit die Bereitschaft zur Nennung von Adressen korreliert ist mit systematischen Verzerrungen in der Population und somit kein repräsentatives Abbild der Grundgesamtheit liefert. Weiterhin ist die Bereitschaft der genannten Netzpersonen zur Teilnahme an der Studie von Interesse. Insofern unterliegen Follow-Up-Interviews einer doppelten möglichen Selektion: a) durch die Bereitschaft der Ego-Personen überhaupt Adressen zu nennen, und b) durch die Teilnahmebereitschaft der genannten Netzpersonen. Von Interesse sind hierbei die Ausschöpfungsquoten, Stichprobeneffekte und die Validität des Namensgenerators hinsichtlich einer nahezu vollständigen Erfassung aller forschungsrelevanten Netzpersonen.

Im zweitgenannten Verfahren der Proxy-Interviews kommt man/frau zwar zu einer billigeren Lösung, allerdings ebenfalls auf Kosten methodischer Probleme. Das Verlassen auf die Informantenangaben von Ego über "seine" Netzpersonen wirft die Frage der Verzerrung und der Fehlperzeption auf. Es gilt zu überprüfen, inwieweit die Angaben von Ego die realen Einstellungen und Merkmale der genannten Netzpersonen wiedergeben oder ob diese Angaben nicht eher auf einer sozialen Projektion beruhen. Zur Beantwortung dieser Frage sind zuverlässige Schätzwerte aus repräsentativen Studien zu egozentrierten Netzwerken notwendig, mithin die Durchführung von Follow-Up-Interviews. Das Argument ist hierbei, daß dadurch das Ausmaß der möglichen Verzerrung bestimmt werden kann. Bestätigen die Follow-Up-Interviews der Netzpersonen die Angaben von Ego über dieselben, ist die Anwendung dieses kostengünstigen Verfahrens zur Erhebung egozentrierter Netzwerke methodisch weitgehend unbedenklich. Werden hingegen gravierende Invaliditäten aufgespürt, ist die Frage evident, ob diese Meßungenauigkeit stabil ist oder variiert. Probleme in der Genauigkeit der Proxy-Daten könnten durch die Zuverlässigkeit der Ego-Angaben überprüft werden. Je zuverlässiger die Angaben von Ego über seine eigenen Merkmale sind, desto zuverlässiger wären auch die Angaben über die Netzpersonen zu unterstellen. Damit stellt sich die Frage des Zusammenhangs von Validität und Reliabilität als einander bedingende Meßkriterien in den Sozialwissenschaften. Stabile Fehlperzeptio-

nen wären im Sinne von Pappi (1977:13f) dahingehend interpretierbar, daß sich Ego in seinem Verhalten gegenüber dieser Netzperson an diesen zwar objektiv falschen Antizipationen orientiert und diese damit subjektiv handlungsleitend sind. Instabile Fehlperzeptionen würden hingegen auf eine willkürliche, eher zufällige Schätzung hinweisen. Zu kontrollieren wäre jeweils, inwieweit Kontaktfrequenz und Regelmäßigkeit mit instabilen und stabilen Fehlperzeptionen verbunden sind und inwieweit die erfragten Informationen über Alter überhaupt bei Ego verfügbar sind.

Insgesamt ist dieses Verfahren weitaus weniger zeit- und kostenaufwendig als das erstgenannte Verfahren der Follow-Up-Interviews. Entsprechend findet sich bei Betrachtung der empirischen Bestandsaufnahme von Netzwerkstudien eine Vielzahl von Studien, in denen egozentrierte Netzwerke alleinig über Ego-Angaben erfaßt wurden. Die methodische Nachfrage nach der Validität dieser Ego-Angaben wird jedoch weitgehend ausgeklammert. Verfügbar sind nur wenige Studien, u.a. die bereits erwähnte Decatur-Studie von Katz/Lazarsfeld (1955, 1965) und die Replikation der Detroit-Area-Studie von Laumann/Pappi (1976). Für eine Vergleichbarkeit unterliegen die Studien starken Restriktionen<sup>1)</sup>. Für die Analyse der Reliabilität der Erhebung von egozentrierten Netzwerken existiert lediglich die Grundlagenstudie von ZUMA (Hoffmeyer-Zlotnik 1987; Pfenning/Pfenning 1987; Pfenning A. 1988). So fällt die methodische Bestandsaufnahme zunächst quantitativ sehr spärlich aus. Es verbleibt die Frage der inhaltlichen Bewertung der Ergebnisse dieser Studien, die auf der Basis empirischer Analysen diskutiert wird.

Katz und Lazarsfeld merken kritisch an, daß lediglich circa 55% der von Ego geäußerten Kommunikationsrollen in der Perzeption von Alter eine Bestätigung finden. Franz U.Pappi und Gunter Wolf weisen in ihrem Beitrag in Meulemann (1984) für die Variable Parteineigung eine systematische Verzerrung nach. Die eigene Parteipräferenz dient demnach als Urteilsanker für die Netzwerkangaben über Dritte. Die (partei)politische Homogenität im sozialen Netzwerk scheint demnach überschätzt zu sein (Pfenning/Pfenning/Mohler 1989). Astrid Pfenning (1988:67) weist für die Variable Parteineigung, bei allerdings ebenfalls statistisch unbefriedigenden Fallzahlen, einen der höchsten relationalen Reliabilitätskoeffizienten (Test-Retest-Korrelation (Cramer's  $V = .83$ )) auf. Dies ist ein Indiz für eine stabile Fehlperzeption. Außerdem ist darauf hinzuweisen, daß Pappi/Wolf die Verfügbarkeit dieser politischen Information bei Ego nicht kontrollierten. Dies wäre jedoch ebenso notwendig wie eine Kontrolle der unterschiedlichen Befragungsformen mündlich versus telefonisch. Insgesamt werfen die Ergebnisse dieser wenigen methodischen Studien eher weitere Probleme auf, als daß sie zur Bestätigung des Informan-

tenverfahrens beitragen. Daß ungeachtet dieser kritischen methodischen Ergebnisse die Grundlagenstudien über egozentrierte Netzwerke nicht weiter verfolgt wurden, um die ausstehenden Fragen der Validität und Reliabilität zu klären, stellt das große methodische Defizit der Netzwerkanalyse dar. Vor diesem Hintergrund wurde im Projektantrag zur vorliegenden Studie über Massenkommunikation und Interpersonale Kommunikation an der Universität Hohenheim der methodische Teil ausdrücklich um die Fragestellung der Validierung von Netzwerkangaben erweitert (Schenk 1988).

## 2. Strategie und Taktik bei der Erhebung egozentrierter Netzwerke

Seit den Arbeiten von Don A. Dillman (1978) über Ausschöpfungsquoten in schriftlichen und telefonischen Meinungsumfragen, ist in den Sozialwissenschaften ein Streit über den richtigen Weg zur Annäherung an den Befragten entbrannt, der einer gewissen Dogmatik nicht entbehrt. Kleinere Studien von Hans-Jürgen Hippler<sup>2)</sup> über den Einfluß von Datenschutzauskünften und eine Vielzahl von Arbeiten amerikanischer Sozialforscher (zusammenfassend Dillman 1979) sind zum Thema Sensibilisierung bei sogenannten heiklen Fragen und zum Thema Datenschutz verfügbar und trotzdem ist keine kausale Antwort möglich. Neben den kaum vergleichbaren Bestimmungen zum Datenschutz im interkulturellen Vergleich und der individuellen Bewertung des Datenschutzes durch unterschiedliche politische Sozialisationen<sup>3)</sup>, wird vor allem der Betroffenheit durch das Befragungsthema zuwenig Beachtung geschenkt. Persönliches Involvement als motivationaler Aspekt zur Teilnahme an Befragungen ist ein den prinzipiellen, themenunabhängigen Vorbehalten gegengerichteter Gesichtspunkt. Insgesamt ist der subjektive Eindruck nicht zu leugnen, daß der einzelne befragte Mensch in der empirischen Sozialforschung zum entpersönlichten Individuum und Datenspende wird. Unter der Voraussetzung, daß jedoch jedes sozialwissenschaftliche Interview auf einer freiwilligen Teilnahme beruht, ist eine Motivation des Befragten zu unterstellen. Diese kann in einer thematischen Betroffenheit oder sozialpsychologischen Beweggründen liegen, wie beispielsweise Neugierde.

Für die Erhebung egozentrierter Netzwerke war unsere Ausgangsposition von folgenden Annahmen und Auffassungen beeinflusst:

a) Die Erhebung von Adressen ist als heikle Frage zu werten und sensibilisiert den Befragten für Belange des Datenschutzes.

b) In dieser Situation ist den motivationalen Aspekten erhöhte Aufmerksamkeit zu schenken, um dieser Sensibilisierung durch Schaffung von Betroffenheit entgegenzuwirken.

c) Unter methodischen Aspekten sollte eine ausführliche Information über Datenschutzbelange und Anonymisierungsmöglichkeiten den Härtegrad für die mögliche Netzausschöpfung darstellen. Soweit die These gilt, daß weniger Datenschutz zu mehr Teilnahmebereitschaft führt, wäre durch unsere Studie somit das untere Limit an Ausschöpfung erreicht.

Zusammenfassend entstand aus diesen Überlegungen ein restriktives Befragungsdesign, das nachfolgend erläutert wird.

### 3. Restriktives Befragungsdesign und Ausschöpfungsquoten

In Replik auf die erwähnte doppelte Selektion der Teilnahmebereitschaft bei egozentrierten Netzwerkstudien mit Follow-Up-Interviews, kommt den Ausschöpfungsquoten eine erhöhte Bedeutung zu. Vorgesehen waren eine mündliche Befragung von circa 900 Personen einer gemeindetypologischen Stichprobe, eine schriftliche Nachbefragung dieser Zielpersonen und eine schriftliche Befragung der genannten Netzpersonen. Das Konzept der gemeindetypologischen Stichprobe basierte auf durch Standortquotienten ermittelte typische Gemeinden in der Region Mittlerer Neckar für die Kategorien dörflich, mittelzentrisch und urban. Es handelte sich um eine Adressenstichprobe aus den Melderegistern. Dies war eine der Bedingungen für die Realisation folgender Maßnahmen zur Erhöhung der Rücklaufquoten und der Teilnahmebereitschaft.

#### 3.1 Bereitschaft zur Adressennennung in der mündlichen Befragung

In Anwendung des eigentlich für schriftliche und Telefonumfragen entwickelten "Total-Design-Method" Verfahrens (TDM) von Dillman wurden alle Zielpersonen, die mündlich interviewt werden sollten (Bruttoansatz von 2400 Adressen), mit einem kurzen informativen Anschreiben über den eventuell bevorstehenden Interviewerbesuch angeschrieben. Zugleich wurden in der lokalen und regionalen Presse der betroffenen Gemeinden bzw. der entsprechenden Region Veröffentlichungen zum Forschungsprojekt platziert. Die Netzpersonen wurden in anonymisierter Form innerhalb des Fragebogens auf einer "Netzkarte" mit Vornamen erfaßt, zunächst ohne Hinweise auf die Adresse oder den vollständigen Namen. Der Namensgenerator bestand aus vier Stimulusvorgaben (wichtige persönliche Dinge besprechen, gesellige Unternehmungen und Aktivitäten, generelle politische Diskussionen und der poli-

tisch informierte Bekannte), basierend auf den Studien von Ronald S. Burt (General Social Survey 1985) und Claude S. Fischer (NCCS-Studie, 1982).

Die Erfassung der Anschriften erfolgte auf einem gesonderten Blatt zu Ende der Erhebung. Den Befragten wurden drei Optionen wahlweise angeboten: Überlassung der Adresse, Überlassung der Telefonnummern mit der Auflage, daß die universitäre Projektgruppe sich mit den betreffenden Personen ins Benehmen setzt, um die Teilnahmebereitschaft zu (hint)erfragen, und Aufschreiben der Adresse auf ein vorgefertigtes Kuvert mit der Versicherung, daß die Adresse von uns nicht aufgenommen wird. Insgesamt waren 46 % (N=406) der Befragten bereit, eine, mehrere oder alle Adressen von Netzpersonen anzugeben. Die Anzahl der von diesen Zielpersonen genannten Netzpersonen beträgt 922 Fälle, davon 892 verwertbare Adressenangaben. Dies entspricht durchschnittlich zwei bis drei Adressen (Mittelwert: 2,3). Dieser Mittelwert ist etwas geringer als die durchschnittliche Netzgröße mit 3,1 Netzpersonen.

### 3.2 Schriftliche Nachbefragung der Zielpersonen und schriftliche Befragung der Netzpersonen

Bei diesen Untersuchungen kam das erwähnte restriktive Erhebungsdesign zur Anwendung. Die Befragten erhielten mit den Voranschreiben und weiteren Erinnerungsschreiben ausführliche Informationen zum Datenschutz, zur Adressenquelle ohne konkrete Nennung der entsprechenden Personen, Hinweise auf die Paginiernummer des Fragebogens, deren Funktion und den Folgen einer möglichen Entfernung sowie eine Datenschutzerklärung mit Darstellung eines anonymisierten Datenfiles.

In Anwendung des TDM-Verfahrens wurden weitere Pressenotizen in den lokalen und regionalen Tageszeitungen plziert sowie einige Rundfunkinterviews durchgeführt. Versandt wurde ein informatives Voranschreiben, anschließend der Fragebogen mit weiterem Informationsmaterial, eine Erinnerungskarte und ein Zweitversand von Fragebogen mit Erinnerungsschreiben. Wissenschaftliche Zielsetzung der Studie war neben den erwähnten methodischen und inhaltlichen Netzwerkfragestellungen die Überprüfung der Agenda-Setting-Funktion der Massenmedien. Die egozentrierten Netzwerke dienten wissenschaftlich zur Operationalisierung des interpersonalen Umfeldes als intervenierenden Faktor auf die Medienwirkungen. Gegenstand war die politische Kommunikation, die zum Zeitpunkt der Studie im öffentlichen, gesellschaftlichen Bereich von der Wiedervereinigungsdebatte und der Währungsunion dominiert wurde. Diese beiden Themen wurden auch in allen Anschreiben als motivationale Hilfe be-



nutzt mit dem Hinweis, daß die Studie auch zur Artikulation der Bürgermeinung zu diesen aktuellen und umstrittenen Themen dienen sollte. Die ausführlichen Datenschutzhinweise bis zur Erklärung der Paginiernummer des Fragebogens begründen die Einstufung als "restriktives Erhebungsdesign", da das Forschungsanliegen sehr transparent dargestellt und der Befragte auf eine mögliche Sensibilisierung hingeführt wurde. Feldphase der mündlichen Befragung war vom 29. Januar bis 16. Februar 1990, für die schriftlichen Befragungen vom 16. März bis Ende Mai 1990.

## 4. Beteiligungs- und Ausschöpfungsquoten

### 4.1 Beteiligungsquoten der postalischen Befragung

Zu Beginn der empirischen Ergebnisse steht die Darstellung der Ausschöpfungsquoten der postalischen Nachbefragung der Zielpersonen und der Befragung der genannten Netzpersonen. Tabelle 1 beinhaltet die entsprechenden Angaben für die postalischen Befragungen innerhalb der Netzwerkstudie der Universität Hohenheim.

Ein Spezifikum des Rücklaufs bei Follow-Up-Interviews von Netzpersonen ist die mögliche Identität von Zielperson und Netzperson-Angabe. Zum einen konnte eine Zielperson als Netzperson einer anderen Zielperson angegeben werden. Dies war in insgesamt 15 Fällen zutreffend. Zum anderen konnte eine Netzperson von mehreren Zielpersonen genannt werden. Diese Mehrfachnennung von Netzpersonen lag in 49 Fällen vor. Nimmt man diese 64 Fälle in die Berechnung der Rücklaufquoten für die Netzpersonen hinzu, ergibt sich sogar eine Beteiligungsquote von 66 % für die postalische Netzpersonen-Erhebung. Diese mehrfach genannten Netzpersonen wurden allerdings nur einmalig befragt und der Datensatz für alle weiteren betroffenen Zielpersonen-Records kopiert. Es ist anzumerken, daß diese Zahl von Identitäts- und Mehrfachnennungen durch das gemeindetypologische Stichprobendesign begünstigt wurde. So resultieren 50 (78 %) der 64 Identitäts- und Mehrfachnennungen aus der Teilstichprobe der dörflichen Stichprobengemeinde. Bei repräsentativ angelegten Studien mit großer räumlicher Streuung und niedrigen Fallzahlen in den einzelnen Sample-Points erscheint dieser Rücklaufeffekt weitgehend vernachlässigbar.

Insgesamt ist die Ausschöpfung der postalischen Befragung der Netzpersonen im Vergleich zu den Rücklaufquoten anderer schriftlicher Umfragen als befriedigend einzuschätzen. Die Teilnahmequote von 62 % bzw. 66 % liegt im obe-

ren Drittel der Ausschöpfung von postalischen Stichproben. Hingegen liegt die Beteiligung der Zielpersonen an der Nachbefragung mit knapp 55% an der unteren Grenze der erwarteten Beteiligungsquote. Nach der optimistischen Hypothese "Je höher der Rücklauf, desto zufriedener der Forscher", ist der Rücklauf der Netzppopulation subjektiv als gut und die Beteiligung der Zielpersonen an der schriftlichen Nachbefragung eher als unbefriedigend einzustufen. Felix Schäfer weist für Beteiligungsquoten eine Korrekturformel für die möglichen Verzerrungseffekte bei verschiedenen Rücklaufquoten aus (Schäfer 1979). Legt man diesen mathematischen Ansatz zugrunde, ist eine Beteiligung von unter 60% als problematisch einzuschätzen.

#### 4.2 Anonymisierungsversuche

Ein interessantes methodisches Detail stellen angesichts des restriktiven Forschungsdesigns die Anonymisierungsversuche durch Entfernen der Fragebogen-Kennziffer dar. Ausdrücklich wurde auf die Bedeutung dieser Kennziffer für die edv-technische Auswertung hingewiesen. Neben den allgemeinen Datenschutzhinweisen und einer gesonderten, ausführlichen Datenschutzerklärung wurde durch das Aufzeigen einer Zahlenkolonne zur edv-technischen Verarbeitung der Daten die Anonymität aufgezeigt. Zur Absicherung bestand eine, für den Befragten unmerkliche Kontrollmöglichkeit zur Reidentifikation der Fragebogenkennziffer. In insgesamt 79 Fällen (7,8% aller Rückantworten aus der Nachbefragung und der Netzpersonenerhebung) mußte auf diese Option zurückgegriffen werden. Für die beiden schriftlichen Erhebungen ergeben sich hinsichtlich dieser Anonymisierungsversuche keine Unterschiede in den relativen Häufigkeiten. Bei der Netzpersonen-Befragung versagte in sechs Fällen allerdings auch diese Option. Inwieweit das restriktive Erhebungsdesign zur Erhöhung der Rücklaufquote beitrug, kann durch die fehlende Vergleichsmöglichkeit nicht eindeutig beantwortet werden. Die angesichts der heiklen Fragestellung insgesamt hohe Rücklaufquote läßt jedoch die Interpretation zu, daß ohne dieses Design und die offene Information ein Teil der Befragten nicht an der Umfrage teilgenommen hätte<sup>4)</sup>.

#### 4.3 Beteiligungseffekte der Erinnerungsmaßnahmen

In Erwägung der problematischen Datenerhebung bei Netzwerkumfragen wurden - wie bereits beschrieben - umfangreiche Erinnerungsmaßnahmen vorgenommen. Diese Maßnahmen erbrachten eine Erhöhung der Rücklaufquote für die postalische Nachbefragung der Zielpersonen von 12,8% und von 6,5% bei der Netzpersonen-Erhebung. Insgesamt sind 9,5% der zurückgesandten Fragebogen Exemplare aus dem Nachversand von Fragebogen im Rahmen der Erin-

nerungsmaßnahmen. Dieser Prozentwert markiert den prozentualen Mindesteffekt der Erinnerungsmaßnahmen. Nicht kontrolliert sind diejenigen Fälle, in denen die Befragten erst auf die Erinnerungsmaßnahmen reagierten, aber für den Rückversand das "alte" Erstexemplar des Fragebogens benutzten.

#### 4.4 Die Ausschöpfung von egozentrierten Netzwerken in Massenumfragen

##### 4.4.1 Ausschöpfungsquoten I: Das Verhältnis von Netzgröße und genannten Netzadressen

In diesem Abschnitt geht es um die Ausschöpfungsanteile bei der Erhebung egozentrierter Netzwerke in Massenumfragen. Die interessierenden Fragestellungen sind, wieviele egozentrierte Netzwerke vollständig vorliegen und inwieweit systematische Verzerrungen durch die doppelte Selektion der Bereitschaft zur Adressennennung und durch die Teilnahme der betreffenden Netzpersonen anzutreffen sind. Tabelle 2 zeigt diese Anteile für die Bezugsgruppe von allen Befragten (Spalte 1) und von allen Befragten mit mindestens einer vorliegenden Rückantwort der genannten Netzpersonen (Spalte 2) auf. Die Anzahl der Zielpersonen mit vorliegender Adressennennung vermindert sich aus Gründen der Nichtlesbarkeit von Adressen und nicht zutreffenden Adressen auf 386 Fälle (44.5%). Von diesem Personenkreis liegt für 261 (70%) Fälle mindestens eine Rückantwort einer Netzperson vor. Der Ausschöpfungsquotient gibt die Relation zwischen Netzgröße und vorliegender Anzahl von Netzrecords an, bezogen auf die Anzahl der Befragten mit verfügbaren Adressen und auf die Anzahl der Befragten, von denen wiederum mindestens eine Rückantwort einer Netzperson vorliegt. In 33,5% aller Fälle mit vorliegenden Rückantworten von Netzpersonen ergibt sich ein Ausschöpfungsquotient von unter 50% für die egozentrierten Netzwerke. In 66,5% dieser Fälle liegt das Verhältnis von Netzgröße und verfügbaren Netzrecords bei über 50%. Dieses beruhigende Resultat wird jedoch wiederum dadurch relativiert, daß lediglich in 25,2% der betreffenden Fallgruppe das vollständige egozentrierte Netzwerk vorliegt.

Logisch anknüpfend an diese erste Darstellung von Ausschöpfungsquoten egozentrierter sozialer Netzwerke ergibt sich die Frage nach dem Zusammenhang von Netzgröße, Ausschöpfung und realisierten vollständigen egozentrierten Netzwerkdaten. Zunächst geht es um das Verhältnis von Netzgröße und genannten Adressen. Basis sind hierfür wiederum alle Befragten mit mindestens einer genannten Alteri-Adresse.

Die Verfügbarkeit bzw. die Bereitschaft zur Adressennennung ist im starken Umfang von der Netzgröße abhängig. Je mehr Personen genannt werden, desto größer ist die Chance, eine Adresse genannt zu bekommen. Ausgenommen die gering besetzten Gruppe mit acht bis zehn Netzpersonen, liegt der Anteil mit mindestens einer verfügbaren Adresse zwischen 57% und 75% relativ gleichmäßig auf hohem Niveau. Allerdings ist, wie Tabelle 1 veranschaulicht, der Anteil der Ausschöpfungsquoten negativ mit der Netzgröße korreliert (Pearsons' $R = -.31$ ). Es gelingt - wie später in Tabelle 5 dargelegt - keine Ausschöpfung vollständiger Netzwerke mit mehr als fünf Netzpersonen.

Nachfolgend wird untersucht, welchen Einfluß soziodemographische Ego-Variablen auf die Bereitschaft zur Adressennennung haben. In diesem Zusammenhang wird unterschieden, inwieweit die Bereitschaft zur Nennung von Adressen von kontextuellen und relationalen Merkmalen abhängig ist. Als kontextuelle Merkmale werden hierbei Wohnort, Geschlecht, Alter, Schulbildung und der sozio-ökonomische Status betrachtet.

Tabelle 4 veranschaulicht, daß die Bereitschaft zur Adressennennung zu Stichprobenverzerrungen führt. Für alle einbezogenen soziodemographischen Variablen finden sich für eine Irrtumswahrscheinlichkeit von 5% signifikante p-Werte. Keiner dieser p-Werte ist allerdings besser als .03, und die Koeffizienten liegen zwischen .08 und .17 auf niedrigem Niveau. Für die Variablen Bildung, Wohnort und insbesondere für die Variable Familienstand sind die statistischen Kennwerte am höchsten. Je kleiner der Wohnort, je höher die Schulbildung sowie wenn man/frau verheiratet ist, desto höher ist im bivariaten Fall die Bereitschaft zur Adressennennung. Für die Variable Familienstand lassen sich lebenszyklische Einflüsse postulieren. Ein Aspekt, der bereits in den Arbeiten von Claude S. Fischer diskutiert wird (Fischer 1977a, 1977b, 1982).

Relationale Merkmale beziehen sich auf die Dyade zwischen Ego und "seinen" jeweiligen Netzpersonen. Für die Bereitschaft zur Adressennennung postulieren wir die Hypothese, daß diese von der Vertrautheit zwischen Ego und der betreffenden Netzperson abhängig ist. Je näher sich beide Personen stehen, desto eher bzw. wahrscheinlicher ist die Adressennennung. Die intervenierende Einflußgröße ist hierbei die Erwartungssicherheit, daß die spontane Nennung der Anschrift gegenüber dem Interviewer nicht auf eine nachträgliche Ablehnung und Sanktion durch die Netzperson stößt. Diese Erwartungssicherheit (im Sinne von Niklas Luhmann in seiner Abhandlung über Vertrauen (1968) zu verstehen), bedarf einer guten Kenntnis und Vertrautheit miteinander. Die meßtechnischen Variablen für diese Dimension der Vertrautheit sind die sub-

jektive emotionale Nähe im Verbund mit der formalen Beziehung, die Bekanntheitsdauer sowie die Kontaktfrequenz. Tabelle 5.1 bis 5.4 stellen die Ausschöpfungsquotienten für diese Variablen dar. Die Ausschöpfungsquotienten bezeichnen hierbei das Verhältnis von realisierten Netzpersonen-Interviews zur Anzahl verfügbarer Adressen je egozentriertem sozialen Netzwerk. Basis sind hierbei alle Befragten, die mindestens eine Adresse einer Netzperson angegeben haben.

Die niedrigen Korrelationskoeffizienten sind eine Folge der durchweg hohen Anteile der hundertprozentigen Ausschöpfungsquoten bei den verschiedenen Prozentquartilen der Anteilswerte der kontextuellen und relationalen Variablen. In allen Spalten weist die Kategorie 100% des Ausschöpfungsquotienten die höchsten relativen Häufigkeitsanteile aus. Zwar ist in der Spalte des höchsten Anteils von Verwandtschaftsbeziehungen, der längsten Bekanntheitsdauer, der regelmäßigsten, häufigsten Kontaktfrequenzen und der intensivsten emotionalen Nähe der relative Anteil egozentrierter Netzwerke mit vollständigem Adressenpool am höchsten, jedoch liegt auch in den anderen Kategorien dieser Anteil sehr hoch und nimmt sogar tentativ bezüglich der Anteilskategorien von 26 bis 99% ab.

#### 4.4.2 Ausschöpfungsquoten II: Das Verhältnis von realisierten Netzwerkinterviews und verfügbarem Adressenpool

Wir haben bisher das Verhältnis von Netzgröße und der Anzahl genannter Adressen betrachtet. In Anschluß an diese Diskussion verbleibt die Frage der Ausschöpfung egozentrierter Netzwerke hinsichtlich der realisierten Interviews. Basis sind hierbei alle Befragten von denen Netzpersonen-Adressen genannt und von denen mindestens eine Rückantwort einer Netzperson vorliegt (N=261). Der Ausschöpfungsquotient der Tabelle 6 gibt entsprechend das Verhältnis zwischen der Anzahl der verfügbaren Netzrecords und der Anzahl der genannten Netzadressen wieder.

In den einzelnen Tabellenspalten sind die Prozentanteile für die jeweilige Bezugsgröße der absoluten Anzahl von genannten Netzadressen eingetragen. Die vollständige Ausschöpfung für die Netzgröße = 1 resultiert aus der Daten-selektion, da nur Befragte mit mindestens einer Rückantwort in die Prozentuierungsbasis eingehen. Die Tabelle zeigt auf, daß lediglich für die Netzgröße bis zu fünf Personen vollständige egozentrierte Netzwerke vorhanden sind. Die Anteile bewegen sich hierfür zwischen 45,8% und 14,8%. Insgesamt ist für 110 Fälle (42% aller Befragten mit Adressenangaben über Netzpersonen) ein vollständiges egozentriertes Netzwerk vorliegend. Jedoch wird

dieses Ergebnis relativiert durch den hohen Anteil von 61 Netzwerken, die nur aus einer Netzperson bestehen.

Wir schließen dieses Kapitel mit dem Vergleich der soziodemographischen Struktur aller genannten Netzpersonen und der Netzpersonen, die an der schriftlichen Befragung teilgenommen haben. Die Daten für alle genannten Netzpersonen entstammen den Angaben von Ego. Es sind Proxy-Daten. Der Vergleich der beiden Befragungsgruppen erfolgt über die prozentualen Anteile für die einzelnen Ausprägungen der Variablen Geschlecht, Alter und Bildung.

Das Ergebnis des Vergleiches der beiden Alteri-Populationen zeigt für die Variable Geschlecht und Alter eher graduelle Unterschiede auf. Die Quote älterer Netzpersonen ist mit 3,9% bis 4,8% in der Rücklaufpopulation höher. Für die Bildungsvariable ergeben sich hingegen deutlichere Differenzen. Der Anteil von Netzpersonen mit niedriger Schulbildung ist in der Rücklaufpopulation deutlich geringer als in der Ausgangspopulation aller von Ego genannten Netzpersonen. Der relative Anteil von Personen mit höherer Bildung ist mit 4,9% (Abitur) bis 6,1% (Hochschule) in der Rücklaufpopulation höher. Es ist ein Bildungseffekt und ein verminderter Alterseffekt auf die Bereitschaft zur Teilnahme an Follow-Up-Interviews zu konstatieren. Bereits in diese Analyse gehen Informantenangaben über Dritte ein. Nachfolgend soll nunmehr dieser Aspekt von egozentrierten Netzwerkstudien beschrieben und diskutiert werden.

## 5. Güte von Proxy-Angaben

Netzwerkanalysen beruhen auf den Angaben von Ego über ausgewählte Merkmale zu Netzpersonen und auf Angaben über die Beziehungen der Netzpersonen untereinander. Nachfolgend werden Analysen über die Zuverlässigkeit und Gültigkeit von Angaben der Zielpersonen über einzelne Netzpersonen erläutert. Der weitere Bereich der Validität von Angaben über die Beziehungen zwischen den einzelnen Netzpersonen ist mit den Daten der Netzwerk-Studie der Universität Hohenheim nicht möglich. Derartige Analysen würden einen weiteren, großen Arbeitsaufwand bedingen und die Beantwortung dieser Fragestellung soll zukünftigen Netzwerkstudien vorbehalten bleiben. Die empirische Bestandsaufnahme zum methodischen Problem der Validität relationaler Merkmale (Pfenning A., 1988) erbringt im Ergebnis den Mangel an verfügbaren Studien. Vergleiche sind zudem beeinträchtigt durch unterschiedliche Namensgeneratoren, unterschiedliche Stichproben und verschiedene Operationalisierungen. Für die Beschränkung auf soziodemographische Variablen ist in

der bundesdeutschen Forschung lediglich die Publikation von Pappi/Wolf (1984) verfügbar. Für die Variablen Alter, Religion, Parteipräferenz und Berufsangabe weisen die Autoren eine Intraklassenkorrelation von maximal .78 und den Anteil prozentualer Übereinstimmung von 88 % (für die Variable Religion) auf.

Für die einzig direkt vergleichbare Variable Alter ist bezüglich der prozentualen Übereinstimmung ein Unterschied von 5 % festzustellen. Allerdings ist in der Korrelation der "Kongruenz-Wert" für die Population der Hohenheimer Studie wesentlich höher, ein Hinweis auf eine deutlich geringere Abweichung der Altersangaben. Für die indirekt vergleichbaren Variablen Berufsangabe und Bildung ergibt sich ein ähnliches Bild im Vergleich beider Studien. Die Variable Geschlecht, für die als invariantes Merkmal eine Korrelation von Eins zu erwarten wäre, ist hier aufgeführt, weil es sich um eine schriftliche Nachbefragung handelt und insofern keine Kontrolle über die den Fragebogen beantwortende Person existierte. Bei einigen Fällen kommt es auch für dieses invariante Merkmal zu Inkonsistenzen, was auf einen Effekt der Befragungsförm schriftlich versus mündlich oder auf Vercodungsfehler zurückzuführen ist. Läßt man diese graduellen Unsicherheiten außer Betracht, zeigen beide Tabellen identische Verteilungen.

Zur Interpretation der in Tabelle 8 und 8.1 dargestellten statistischen Kennwerte bedarf es theoretischer und methodischer Annahmen. Da es sich um invariante und zudem retrospektive Daten handelt, läßt sich eine sehr hohe bis vollkommene Übereinstimmung erwarten. Die Daten-Attribuierung "invariant" bedeutet, daß keine Veränderungen des Wertes möglich sind und somit ein wahrer Wert vorhanden ist. Retrospektive Daten sind in diesem Sinne invariant, weil sie sich auf vergangene Ereignisse beziehen. Diese Einstufung läßt die ausgewählten soziodemographischen Variablen als Härtestest für die tatsächliche Kenntnis von Ego über die betreffende Netzperson erscheinen. Der intervenierende Faktor Kenntnis bzw. Informationsverfügung weist ein Meßkontinuum von genauer Wiedergabe erhaltener Informationen bis zur "ego-eigenen" Schätzung und Interpolation des Alteri-Merkmals aus anderen Variablen auf. Es ist eine lineare Funktion zu unterstellen: Die Wiedergabe von Informationen über Alter, die Ego von demselben erhalten hat, sollte eine höhere Genauigkeit aufweisen als die von Ego ohne direkte Alter-Information vorgenommene Schätzung und Projektion. Für die Meßgenauigkeit der Schätzung ist wiederum die Variationsbreite des betreffenden Merkmals bedeutsam. Je weniger Ausprägungen die Variable aufweist, desto höher ist die Gefahr einer Fehlschätzung gegeben. Im dichotomen Fall kann die Ego-Schätzung nur

falsch oder wahr sein (Beispiel Geschlecht). Bei großer Variationsbreite eines Merkmals steigt die Wahrscheinlichkeit einer zufällig annähernden Schätzgenauigkeit. Auf der Basis dieser Annahmen sind die Ergebnisse zur Genauigkeit der Ego-Angaben mit einem Indikator zur Informationsverfügung zu überprüfen. Hierfür bieten sich die relationalen Merkmale Bekanntheitsdauer, Kontaktfrequenz und affektive soziale Intensität der Dyade an. Als weiteren Indikator zur Unterscheidung von Schätzungen und Informationswiedergabe verwenden wir als Hilfskonstruktion die Eigenreliabilität des Befragten über eigene Merkmale. Diese Werte stehen durch die Nachbefragung der Zielpersonen zur Verfügung.

Tabelle 9 weist auch für die Eigenreliabilität keine vollständige Kongruenz auf, obwohl es sich um retrospektive und invariante Angaben handelt. Die Abweichung bei der Variable Geschlecht beruht auf einem Instrumenteneffekt von mündlicher und schriftlicher Erhebung. Bei einigen Zielpersonen der mündlichen Befragung, die den Lebenspartner(in) als Netzperson nannten, wurde der schriftliche Fragebogen von diesen Lebenspartnern ausgefüllt. Dieser Effekt von mündlicher (Vor)Befragung und schriftlicher Nachbefragung bei Erfassung egozentrierter Netzwerke trat in 39 Fällen auf. Er wurde im Design der Studie ursprünglich nicht berücksichtigt. Die hohe Anzahl von Fehlperzeptionen bedingt den Schluß, daß bei Durchführung von Follow-Up-Interviews mit Partnernennungen erhöhte Sorgfalt auf die Adressierung der erwünschten Zielperson zu verwenden ist. Bereinigt man den Datensatz um diese Fälle, verändern sich die Test-Retestkorrelationen wie in Tabelle 9.1 dargestellt.

Geschlecht und Alter weisen nunmehr eine annähernde vollständige Kongruenz auf. Die Variablen Bildung und Wahlabsicht für die damals bevorstehende Bundestagswahl 1990 weisen in der Test-Retestanalyse die niedrigsten Koeffizienten auf. Allerdings ist die Variable Wahlabsicht kein invariantes Merkmal und insoweit sind wahre Einstellungsveränderungen - zumal in der Zeit heftiger Kontroversen um die Themen Wiedervereinigung und Währungsunion - möglich. Der niedrige Test-Retest-Koeffizient ist insofern nicht alleinig auf statistische Inkonsistenzen, sondern auch auf inhaltliche Gründe zurückzuführen und zu erklären.

### 5.1 Der Einfluß von Kontexteffekten auf die Güte von Proxy-Daten über Netzpersonen

Der Einfluß von Kontexteffekten bedingt eine genügende Streuung in den Merkmalen. Diese Voraussetzung ist für Netzwerkanalysen umso mehr zu prüfen, da die Erhebung der Netzpersonen sich oftmals auf das primäre soziale Milieu des Befragten bezieht. Dieser - eigentlich unerwünschte - Gruppen-



effekt (Schenk 1983, 1984) bedingt Meßartefakte für kontextuelle Variablen. In der Netzwerkstudie der Universität Hohenheim wurde deshalb versucht die Meßgüte des Namensgenerators dahingehend zu verbessern, daß genügend Streuung in den Kontextmerkmalen zustandekommt. Die Darstellung des Erhebungsverfahrens ist nicht Thema dieser Publikation, weshalb nachfolgend nur die Häufigkeitsverteilungen dargestellt werden<sup>5)</sup>. Als Gestaltmaß ist die Schiefe aufgeführt, deren Kennwerte für alle Variablen - am deutlichsten für die Variable Bekanntheitsdauer - eine rechtsschiefe Verteilung aufzeigen. Jedoch erreicht keiner dieser Kennwerte die kritische Marke eines z-Wertes für eine Irrtumswahrscheinlichkeit von 10% ( $Z=1,64$ ). Damit ist die Annahme haltbar, daß die Variablen eine (annähernde) Normalverteilung aufweisen.

Auf dieser statistischen Basis erfolgt die Analyse möglicher Kontexteffekte über bivariate Korrelationen. Geprüft wird, ob die abhängige Variable Übereinstimmung versus Nicht-Übereinstimmung von bestimmten Ausprägungen der Kontextmerkmale überzufällig beeinflusst wird. Es finden sich keine deutlichen signifikanten Unterschiede zwischen den untersuchten Kontextvariablen und der Übereinstimmung der Proxy-Angaben über Netzpersonen. Die Tendenz einiger Befragter zu unrelia­blen und invaliden Antworten steht in keinem Zusammenhang mit Variablen, die über Zeitdauer, Kontaktfrequenz und affektive Nähe der Beziehung indirekt die verfügbaren Informationsressourcen von Ego erfassen sollten. Nach Ausscheiden dieser Kontextvariablen als Determinanten der Meßgenauigkeit von Proxy-Interviews, stellt sich die Frage der Bedeutung individueller Merkmale.

## 5.2 Ego-Reliabilität und Netzpersonen-Proxys im Zusammenhang

Die Reliabilität der Eigenangaben von Ego soll in dieser Analyse als Schätzparameter für die Genauigkeit, das Maß der Übereinstimmung von Proxy-Angabe und Angabe der Netzpersonen dienen. Das Erhebungsdesign sah aus pragmatischen Gründen keine Nachbefragung der Netzpersonen vor, die zusätzlich zur Kontrolle der Netzpersonen-Angaben sinnvoll erscheint. Die nachfolgenden Analysen basieren deshalb auf der Annahme, daß die Angabe der Netzperson als der "wahre Wert" betrachtet wird. Die Intervention der möglichen Unzuverlässigkeit der Netzpersonen-Daten bleibt also außer Betracht, obwohl diese gleichermaßen wie die Ego-Eigenangaben wahrscheinlich keine vollkommene Reliabilität aufweisen. Die pragmatische Begründung liegt in der erwarteten Reduktion der Fallzahl. Die Auswertung bezieht sich auf diejenigen Fälle, in denen die Zielperson an beiden Erhebungen und zudem mindestens eine Netzperson an der schriftlichen Befragung teilgenommen hat. Dies sind 141 Fälle. Verteilt auf die Relationen der Netzpersonen ergeben sich 391 Dyaden. Die

Analyse beruht auf diesem relationalen Datensatz, der durch die Duplizierung der Ego-Angaben auf die entsprechenden Netzrelationen generiert wurde<sup>5)</sup>. Eingegangen in diesen Datensatz sind die Variablen Alter (kontinuierliche Ausprägung nach Lebensjahren) und Bildung (vierstufig: Volksschule, Realschule, Abitur, Universitätsabschluß) von Ego und den Netzpersonen. Die Variable Geschlecht bleibt unbeachtet, da sie nach den Vorkontrollen und Datenbereinigungen eine vollständige Kongruenz der Angaben aufweist. Die Auswertung basiert auf einem Paarvergleich der Ego-Reliabilität (Variablenpaar 1) mit der Güte der Proxy-Daten (Variablenpaar 2). Die sich ergebenden vier Kombinationsmöglichkeiten sind mit den jeweiligen relativen Häufigkeiten für die Variablen Alter und Bildung in Tabelle 11 dargestellt.

Für die Variable Bildung sind 55 %, für die Variable Alter lediglich 26 % aller Angaben der Zielperson über sich und zu den Netzpersonen-Proxys identisch. Für die Variable Alter gilt hierbei, daß auf Übereinstimmung in den Lebensjahren getestet wurde. Da auch keine geringfügigen und durchaus logisch möglichen Abweichungen (Geburtsjahr zwischen Erst- und Zweitbefragung) berücksichtigt wurden, kann dieses Vorgehen als konservativer Test bezeichnet werden. Die Angaben über identische Proxys zu den Netzpersonen liegen bei Prozentwerten von 30,7 % für die Variable Alter und 67 % für die Variable Schulbildung. Insgesamt sind 16 % aller Angaben bei der Variablen Alter und 19 % bei dem Merkmal Bildung sowohl unreliabel als auch invalide. Diese Anteile charakterisieren gemeinsam mit denjenigen Fällen, in denen die Eigenangaben nicht kongruent sind, wohl aber die Proxys, die kritische Population.

Im nächsten Schritt galt es, den Zusammenhang von Reliabilität und Proxy-Validität für beide Variablen zu untersuchen. Unter Beibehaltung der vierstufigen Kombination für jede Variable ergeben sich nunmehr 16 Kombinationsmöglichkeiten.

In 17,1 % aller Fälle stimmen die Eigenangaben und die Netzpersonen-Proxys vollkommen überein. 66,9 % der Fälle weisen eine Übereinstimmung der Eigenangaben für beide Variablen auf und in 10,6 % aller Fälle findet sich keine Kongruenz der Eigenangaben für beide Variablen. Keine Reliabilität und keine Validität in einer oder beiden Variablen für die Eigenangaben und für die Netzproxys sind in 22,1 % aller Fälle festzustellen, davon 3,9 % die vollkommen invalide und unreliabel sind. Bei 67 % findet sich eine Inkongruenz in den Proxy-Angaben für mindestens eine der analysierten Variablen. Analog zur Vorgehensweise in Abschnitt 5.1 wurde der Zusammenhang von Reliabilität und Validität kontrolliert auf mögliche kontextuelle Effekte. Für die formale Beziehung (Spearman-Rangkorrelation = .04), Bekanntheitsdauer (Spearman-

Rangkorrelation= .09) und Kontaktfrequenz (Spearman-Rangkorrelation= .02) sowie die affektive Nähe (Spearman-Rangkorrelation= .03). Es ergeben sich keine signifikanten Kontexteffekte.

Betrachten wir die Test-Retest-Korrelationen für die Fallgruppe der Zielpersonen ( $N=141$ ), die an beiden Befragungen teilgenommen haben und für die zudem eine Rückantwort einer Netzperson vorliegt ( $N=391$  Relationen). Im Vergleich mit den Eigenwerten für alle Zielpersonen, die in Tabelle 9 dargestellt sind, sind höhere Test-Retest-Korrelationen festzustellen. Dies gilt auch für den Vergleich mit den Proxy-Analysen für alle Netzpersonen (vgl. Tabelle 8.1). Nach diesen Befunden scheint die Teilnahmebereitschaft an Befragungen eine intervenierende Größe für die Zuverlässigkeit von Informationen zu sein. Je ausgeprägter die Teilnahmebereitschaft an den Befragungen ist, desto besser ist die Reliabilität der Eigenangaben und die Validität der Proxy-Daten. Die hohe Korrelation für die Variable Alter zeigt auf, daß nur geringe Streuungen zwischen den beiden Messungen vorhanden sind. Diese führen jedoch bei konservativen Tests auf vollständige Kongruenz zu den geringen Übereinstimmungsquoten im Paarvergleich.

Die bisherigen Analysen basieren - insbesondere für die kontinuierliche Variable Alter - auf der konservativen Voraussetzung vollkommener Datenidentität. Begründet wird dies durch die Tatsache, daß es sich wie beschrieben um invariante Merkmale handelt. Zur erweiterten Interpretation der Güte von Proxy-Daten werden nachfolgend relationale Korrelationsmaße berechnet, die auch das Ausmaß der Streuung für die Variable Alter einbeziehen.

Der Anteil kongruenter Eigenangaben und Proxy-Daten nimmt um etwa 26 % zu. Insgesamt sind nunmehr circa 43 % vollkommen valider Fälle anzutreffen. Weitere Kenndaten dieser Tabelle sind: In 72,9 % aller Fälle sind die Eigenangaben kongruent, in 58,4 % trifft dies für die Proxy-Angaben zu, die Anteile inkongruenter, das heißt unreliabler und invalider Fälle werden marginal. Die Korrelationsmaße sind im Vergleich zu Tabelle 12 allerdings fast unverändert. Weiterhin ein Indiz dafür, daß starke Abweichungen einzelner Fälle vorhanden sind. Die Schätzung der Validität der Proxy-Angaben über Netzpersonen aus den Eigenangaben der Befragten für die soziodemographischen Variablen ist nach den bisherigen bivariaten Analysen auf einer hinreichenden statistischen Basis gewährleistet. Die Inkongruenzen sind weitgehend auf Abweichungen der Proxy-Angaben zurückzuführen. Und diese sind nach prozentualen Anteilen durchaus gewichtig.

## 6. Fazite und Defizite

Inhalt dieses Beitrages waren die Beteiligungs- und Ausschöpfungsquoten von egozentrierten Netzwerkstudien, die den sozialwissenschaftlichen Standard an Massenumfragen entsprechen. Ergänzt wurde diese Fragestellung durch die Einbeziehung der methodischen Überprüfung der Güte von Informantendaten. Hinsichtlich der forschungspraktischen Effizienz von Follow-Up-Interviews, das heißt den Beteiligungs- und Ausschöpfungsquoten, finden sich befriedigende Resultate. 44 % der Befragten sind zur Angabe aller oder einiger Adressen ihrer genannten Netzpersonen bereit. Für diese Bereitschaft zur Adressenweitergabe lassen sich keine übermäßigen signifikanten Einflüsse soziodemographischer Ego-Merkmale finden und insofern ist per se keine Verzerrung der Teilnahmebereitschaft der Zielpersonen festzustellen. Auch die Beteiligung der Netzpersonen mit 62 % bzw. 68 % ist keinesfalls unterdurchschnittlich. Die Anwendung eines restriktiven Erhebungsdesigns nach den Kriterien Offenheit, weitgehende Informationen über das Projekt und zum Datenschutz sowie mehrfache Kontaktierung hat sich bewährt. Stuft man die Erhebung von Adressen als heiklen Erhebungsbereich ein, sind diese Beteiligungsquoten im Vergleich zu anderen Studien über "heikle Fragen" sogar als außerordentlich hoch einzuschätzen. Trotz der hohen Beteiligung von Netzpersonen sind die Ausschöpfungsquoten an vollständig vorliegenden egozentrierten Netzwerken gering. Vielmehr liegen von vielen egozentrierten Netzwerken jeweils einige Angaben über Netzpersonen vor. Die Fallzahl der Ego-Population vermindert sich hinsichtlich der Bedingung zumindest eines vorliegenden Netzrecords auf 30 %. Der Anteil vollständiger Netzwerke liegt lediglich bei etwa 8 % der ursprünglichen Ego-Population und betrifft ausschließlich Netzwerke kleineren Umfangs.

Unser erstes Fazit ist, daß Follow-Up-Interviews auf der Basis zumindest durchschnittlicher Beteiligungsquoten ohne größere Verzerrungen in der Ego-Population realisierbar sind. Die variierenden Ausschöpfungsquoten im Verbund mit der geringen Ausschöpfung vollständiger Netze bedingen allerdings einen großen Bruttoansatz von realisierten Ego-Interviews für ausreichende Zellenbesetzungen bei Netzwerkanalysen. Das zweite Fazit bezieht sich auf stichprobentheoretische Implikationen bei der Rücklaufpopulation der Netzpersonen. Hierbei finden sich bedeutsame Bildungseffekte und abgeschwächt ein Alterseffekt. Ältere Netzpersonen und Netzpersonen mit hoher Schulbildung haben sich überdurchschnittlich an der schriftlichen Befragung der Netzpersonen beteiligt. Demnach ist auf der zweiten Stufe der Durchführung von Follow-Up-Interviews eine Verzerrung von Ausgangspopulation und Rück-

laufpopulation der Netzpersonen zu erwarten. Diese Verzerrung ergibt sich primär aus den soziodemographischen Merkmalen der Netzpersonen und nicht aus relationalen Merkmalen wie beispielsweise der Bekanntheitsdauer, der Intensität und Kontaktfrequenz oder kontextuellen Merkmalen. Das dritte Fazit zur Frage der Güte von soziodemographischen Proxy-Daten regt zur weiteren Diskussion an. Validität und Reliabilität sind im Paarvergleich der beiden soziodemographischen Variablen nur auf schwacher Basis miteinander korreliert: Nur 18 % aller Dyaden erfüllen diese methodischen Gütekriterien vollkommen, wenngleich nur 3,8 % zum anderen Pol vollkommener Nichtübereinstimmung zählen. Für die Variable Bildung beläuft sich der Anteil kongruenter Daten auf rund 55 %, bei der Variablen Alter auf 26 %. Die Überprüfung der Kongruenz der Variablen Alter nach Lebensjahren ohne Zulassung einer Abweichung bedingt diese prozentualen Unterschiede. Läßt man/frau eine Abweichung von plus/minus einem Jahr zu, ergeben sich hinsichtlich des Anteils kongruenter Daten deutlich verbesserte Werte: 43 % vollkommene Kongruenz von reliablen Eigenangaben und validen Proxy-Angaben und nur marginale Anteile für inkongruente Fälle.

In der vergleichenden Bewertung der Eigenwerte und der Proxy-Daten zeigt sich ohne zulässige Abweichung ein diametrales Resultat. Zwei Drittel der Befragten nennen identische Eigenwerte für beide soziodemographischen Variablen, zugleich wird in 67 % eine ungenaue Angabe zu den Proxy-Daten geäußert. Der geringe Anteil von 3,9 % für die vollkommene Inkongruenz von Eigen- und Proxy-Angaben darf nicht über diesen hohen Anteil von Inkonsistenzen in den wechselnden Kombinationen beider Variablen hinwegtäuschen. Die Korrelationswerte für die Test-Retestanalysen zeigen auf, daß diese Abweichungen auf geringen Streuungen beruhen. Bei der beschriebenen, zulässigen Abweichung für die Variable Alter gleichen sich die Anteile kongruenter Nennungen von Reliabilität und Validität an. Für nominal- und ordinalskalierte Daten bedeuten allerdings auch die verbleibenden Abweichungen gravierende inhaltliche Inkonsistenzen.

Diese Ergebnisse weisen in die Richtung einer kritischen Beurteilung von Angaben über Dritte. Das Zustandekommen von Unreliabilitäten und Invaliditäten ist weitgehend unbeeinflußt von kontextuellen Effekten und relationalen Merkmalen. Dieses Ergebnis weist auf die Bedeutung invarianter Merkmale hin. Bildung und Alter müssen als Informationen lediglich einmal verfügbar sein, da sie als retrospektive Merkmale keinen Veränderungen unterliegen. Zudem ist denkbar, daß diese Informationen zum Standard-Repertoire sozialer Beziehungen zählen, etwa hinsichtlich der Homophilie-Aspekte und Kriterien

bei der Aufnahme sozialer Beziehungen. Die geringe Relevanz kontextueller Variablen für die Güte soziodemographischer Proxys verweist ebenso auf die Bedeutung individueller Merkmale. Gilt die Annahme, daß die Zuverlässigkeit und Validität relationaler Angaben von individuellen Merkmalen abhängig sind, müßten sich die egozentrierten Netzwerke nach den determinierenden Ego-Variablen unterscheiden lassen. Unreliable und invalide Ego-Angaben wären demnach hoch korreliert mit dem Anteil invalider Messungen im egozentrierten Netzwerk. Dieses Defizit der Analyse ist in diesem Beitrag nicht aufgeführt, kann aber mit den Daten der Netzwerkstudie der Universität Hohenheim geleistet werden. Die bisherigen Resultate der relationalen Analysen über Dyadenmerkmale und Ego-Variablen konnten lediglich die geringe Bedeutung der kontextuellen Merkmale und der relationalen Merkmale wie Bekanntheitsdauer, Kontaktfrequenz, affektive Nähe und Beziehungsart zwischen Ego und Alter aufzeigen. Die (ausstehende) Erweiterung der Analyse unter Hervorheben der Strukturmerkmale (z.B. Geschlossenheit, Dichte, Heterogenitäts- und Homogenitätsmaße) und individueller Ego-Merkmale ist ein Beleg für die eingangs formulierte Notwendigkeit "echter" Netzwerkanalysen. Dies bedeutet die gleichzeitige Einbeziehung individueller Merkmale, relationaler Variablen und von Netzstrukturmerkmalen und deren analytische Gewichtung zueinander. Die Einbeziehung von Einstellungen wurde in dieser Abhandlung ebenso außer Betracht gelassen. Da persönliche Einstellungen einem Wandel unterliegen können und somit die Annahme der klassischen Meßfehlertheorie über einen wahren Wert fraglich wird, sind diese Analysen gesondert zu behandeln. Ebenso ausstehend ist die Analyse über die Bedeutung von Projektionen: der Verzerrung der Proxy-Angaben hinsichtlich der Merkmale der Zielperson. Für diese Analyse sind die Fälle mit kongruenten Eigenangaben und inkongruenten Proxy-Angaben interessant. Diese Analyse bleibt als "externes Defizit" in diesem Beitrag ebenfalls ausgeklammert.

In einer zusammenfassenden Bewertung der hier untersuchten Beteiligungs- und Ausschöpfungsquoten, der Berücksichtigung der möglichen Verzerrungen, empirisch genügender Ausschöpfungen und dem Ausmaß der Gültigkeit von Proxy-Daten sind Netzwerkerhebungen in Massenumfragen mit Follow-Up-Interviews möglich. Der Anteil invalider Proxy-Daten ist zwar beträchtlich, kann aber nach den vorliegenden Resultaten aus Strukturmerkmalen und der Reliabilität der Ego-Angaben hinreichend geschätzt werden. Dies bedeutet, daß Follow-Up-Interviews für Subgruppenanalysen interessante Inkonsistenzen aufdecken, für Gesamtanalysen aber nicht zwingend notwendig sind. Unter Gesamtanalysen wird hierbei eine Analyse ohne Spezifizierung von Subgruppen verstanden. Die Inkonsistenzen bei invarianten soziodemographischen Va-

riablen liegt für Proxy-Daten bei circa 42%. Die Irrtumsbandbreite ist bei der überwiegenden Zahl von Befragten zudem sehr gering. Es gilt zudem, daß unreliable Eigenangaben auch zu invaliden Proxy-Daten führen, ebenso wie reliable Angaben überwiegend mit validen Proxy-Daten verbunden sind.

## Anmerkungen

- 1) Die Restriktionen sind in den unterschiedlichen Namensgeneratoren, im Stichprobendesign (in der Detroit-Area-Study werden beispielsweise nur Männer der "weißen" Bevölkerungspopulation befragt) und den geringen Fallzahlen zu sehen.
- 2) Hans-Jürgen Hippler et al. (1990) untersucht anhand einer kleinen Gemeindestichprobe die Teilnahmebereitschaft an einer postalischen Befragung in Abhängigkeit von der Ausführlichkeit beigefügter Datenschutzhinweise. Das Ergebnis läßt sich wie folgt zusammenfassen: Je weniger Datenschutz, desto höher der Rücklauf.
- 3) Für den Einfluß des Datenschutzes auf die Teilnahmebereitschaft lassen sich für den bundesdeutschen Raum vor allem historische und mediale Einflüsse vermuten. Die totale Erfassung und Bepitzelung im Verbund mit dem Denunziantentum während der Nazi-Zeit hat für dieses Thema Sensibilisierungen hinterlassen. Durch verschiedene Ereignisse in den vergangenen Jahrzehnten (Volkszählung, Quellensteuer, Staatssicherheitsdienst in der Ex-DDR) ist das Thema Datenschutz im Blickpunkt der öffentlichen Diskussion geblieben und aktualisiert worden.
- 4) Im ZUMA-Projekt "Heikle Fragen" wird - bei allerdings von unserem Forschungsprojekt deutlich verschiedenen Fragestellungen - über illegale Verhaltensweisen ein Rücklauf zwischen 25 bis 30% berichtet. Die Kontrollmöglichkeit in unserer Untersuchung bezog sich lediglich auf die bereits anonymisierten Identifikationsvariablen, nicht auf Personen bzw. Adressen. Deshalb konnte in einigen Fällen auch keine Zuordnung mehr getroffen werden.
- 5) Diese Datenanforderung kann optimal durch ein Datenbanksystem gelöst werden. Verwendet wurde das Programm SIR/DBMS, das eine Schnittstelle zu den gängigen Statistikprogrammen beinhaltet.

## Literatur

- Burt, Ronald S., 1984: Network-Items should be included in the General Social Survey. Department for Sociology and Center for the Social Sciences pre-print-series. ColumbiaUniversity, New York (mimeo).
- Dillman, Don A., 1978: Mail and Telephone Surveys. The Total Design Method, New York.
- Fischer, Claude Simon, 1977a: The Contexts of Personal Relations.
- Fischer, Claude Simon et al., 1977b: Network and Places: Social Relations in the Urban Setting. New York: The Free Press.
- Fischer, Claude Simon, 1982: To Dwell Among Friends. Personal Networks in Towns and City. Chicago.
- Hippler, Hans-Jürgen/Schwarz, Norbert/Singer, Eleanor 1990: Der Einfluß von Datenschutzzusagen auf die Teilnahmebereitschaft in Umfragen, ZUMA-Nachrichten 27: 54-67.
- Hoffmeyer-Zlotnik, Jürgen, 1987: Egozentrierte Netzwerke in Massenumfragen 1: Zum Design des Methodenforschungsprojektes, ZUMA-Nachrichten 20: 37-43.
- Katz, Elihu/Lazarsfeld, Paul-Friedrich, 1955 (2.te Auflage 1965): Personal Influence, Chicago.
- Laumann, Edward/Pappi, Franz-Urban, 1976: Networks of Collective Action: A Perspective on Community Influence Systems, New York.
- Luhmann, Niklas 1968: Vertrauen, Enke-Verlag, Stuttgart.
- Pappi, Franz-Urban/Wolf, Gunter, 1984: Wahrnehmung und Realität sozialer Netzwerke. Zuverlässigkeit und Gültigkeit der Angaben über beste Freunde im Interview, S. 281-300 in: Meulemann, R./Reuband, K.-H. (Hrsg.): Soziale Realität im Interview - Empirische Analysen methodischer Probleme, Frankfurt.
- Pappi, Franz Urban (Hrsg.), 1987: Methoden der empirischen Sozialforschung, Band 1: Netzwerkanalyse, Tübingen.
- Mitchell, John Clyde, 1969: Social Networks in Urban Situations. Analysis of Personal Relationships in Central African Towns, Manchester, Groß-Britannien.
- Mohler, Ph. Peter/Pfenning, Uwe, 1987: Egozentrierte Netzwerke in Massenumfragen 3: Datenorganisation in einer SIR-Datenbank, ZUMA-Nachrichten 20: 1987: 51-56.
- Pfenning, Astrid, 1988: Zur Zuverlässigkeit von Netzwerkangaben - Empirische Analysen anhand der ZUMA-Methodenstudie "Egozentrierte Netzwerke in Massenumfragen", Diplomarbeit an der Universität Mannheim.
- Pfenning, Astrid/Pfenning, Uwe, 1987: Verschiedene Instrumente - verschiedene Ergebnisse?, ZUMA-Nachrichten 21: 64-77.
- Pfenning, Astrid/Pfenning, Uwe/Mohler, Peter Ph., 1989: Parteipräferenzen in sozialen Netzwerken, ZUMA-Nachrichten 24: 73-86.



- 
- Pfenning, Astrid/Pfenning, Uwe/Mohler, Peter Ph., 1991: Zur Reliabilität von egozentrierten Netzwerken in Massenumfragen, ZUMA-Nachrichten 28: 92-109.
- Porst, Rolf, 1991: Technischer Bericht Nr. 21/1992 zum ZUMA-Projekt "Heikle Fragen", ZUMA, Mannheim.
- Schäfer, Felix, 1979: Musterstichproben-Pläne. Hrsg. vom Arbeitskreis Deutscher Marktforschungsinstitute (ADM), Verlag Moderne Industrie, München.
- Schenk, Michael, 1983: Das Konzept des sozialen Netzwerkes, Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 25: 88-104. Westdeutscher Verlag, Opladen.
- Schenk, Michael, 1984: Soziale Netzwerke und Kommunikation. Mohr-Verlag, Tübingen.
- Schenk, Michael, 1988: Massenkommunikation. Opladen.
- Schenk, Michael, 1984: Soziale Netzwerke und Kommunikation, Mohr-Verlag, Tübingen.
- Schenk, Michael, 1988: Massenkommunikation und interpersonale Kommunikation - Zur Rolle egozentrierter Netzwerke, Forschungsantrag an die Deutsche Forschungsgemeinschaft (DFG), Universität Hohenheim, Stuttgart.

**Tabelle 1: Beteiligungs- und Ausschöpfungsquoten der postalischen Erhebungen bei den Zielpersonen und den Netzpersonen**

| Bezugsgruppe            | Ziel-<br>personen | Netz-<br>personen | Summe |
|-------------------------|-------------------|-------------------|-------|
| Angeschriebene Personen | 899               | 922               | 1821  |
| Neutrale Ausfälle       | 14                | 30                | 44    |
| Prozentuierungsbasis    | 885               | 892               | 1777  |
| Rücklauf                |                   |                   |       |
| absolut                 | 485               | 550               | 1035  |
| in %                    | 54.8              | 62.0              | 58.2  |

**Tabelle 2: Ausschöpfung von Netzwerkdaten**

| Basis   | Alle befragte<br>Zielpersonen |      | Befragte mit<br>Adressen | Befragte mit<br>Rückantwort |
|---|-------------------------------|------|--------------------------|-----------------------------|
|   | abs.                          | in % | in %                     | in %                        |
| Anzahl von Ego-Personen mit<br>verwertbarer Adressennennung | 387                           | 44.5 | 100                      |                             |
| Anzahl von Ego-Personen mit<br>mindestens einem Netzrecord* | 261                           | 30.0 | 70.0                     | 100                         |
| Ausschöpfungsquotient                                       |                               |      |                          |                             |
| keine (0 %)   | 593                           | 69.0 | 28.0                     |                             |
| 0.01 bis 50%  | 90                            | 10.5 | 24.1                     | 33.5                        |
| 50.01 bis 60%   | 57                            | 6.7  | 15.5                     | 21.6                        |
| 60.01 bis 70%   | 29                            | 3.4  | 8.3                      | 11.6                        |
| 70.01 bis 80%   | 9                             | 1.2  | 3.1                      | 4.3                         |
| 80.01 bis 90%   | 8                             | 1.2  | 2.8                      | 3.9                         |
| 90.01 bis 99%   | -                             | -    | -                        | -                           |
| alle (100%)   | 68                            | 8.0  | 18.1                     | 25.2                        |

\* = Netzrecord ist die datentechnische Bezeichnung für einen vorliegenden Datensatz einer Netzperson

Tabelle 3: Netzgröße und Anzahl genannter Adressen (N=387)

| Anzahl von<br>Netzen mit.... | Anzahl von Zielpersonen mit<br>einer oder mehreren Adressenangaben |      |      |
|------------------------------|--|------|------|
|                              | abs.   | abs. | in % |
| 1 Netzperson                 | 86   | 62   | 71.8 |
| 2 Netzpersonen               | 81   | 60   | 72.8 |
| 3 Netzpersonen               | 83   | 57   | 67.5 |
| 4 Netzpersonen               | 51   | 29   | 56.9 |
| 5 Netzpersonen               | 39   | 27   | 69.2 |
| 6 Netzpersonen               | 20   | 13   | 65.0 |
| 7 Netzpersonen               | 12   | 9    | 75.0 |
| 8 Netzpersonen               | 3  | 2    | 66.7 |
| 9 Netzpersonen               | 4  | 1    | 25.0 |
| 10 Netzpersonen              | 8  | 1    | 12.5 |

Chi-Quadrat=120.24; Pearsons' R=.763;  $p < .00001$

Tabelle 4: Bivariate Zusammenhänge von soziodemographischen Ego-Merkmalen und Adressennennung von Netzpersonen.

| Variable      | Chi-Quadrat | Cramers' V/Phi | Koeffizient | p-Wert |
|---------------|-------------|----------------|-------------|--------|
| Geschlecht    | 6.75        | .086           | -           | .037   |
| Alter         | 5.24        | .073           | .063*       | .046   |
| Wohnort       | 24.31       | .164           | .159**      | .033   |
| Bildung       | 19.12       | .146           | .080**      | .033   |
| Ökon. Status  | 8.83        | .099           | .092*       | .033   |
| Familienstand | 27.05       | .173           | -           | .031   |

Für folgende Variablen liegt eine Recodierung in ordinalskalierte Gruppen vor: Wohnort: urban, mittelzentrisch, dörflich; Bildung und ökonomischer Status (gemessen als Sozialprestige-Index nach Mayer-Score-Werten): niedrig, mittel, hoch. Es handelt sich nicht um adjustierte p-Werte.

\* = gemessen mit Pearsons' R; \*\* = gemessen mit Spearmans's Rangkorrelationsmaß

**Tabelle 5.1: Ausschöpfungsquotient und Verwandtschaftsanteile im egozentrierten Netzwerk (Spaltenprozente)**

| Ausschöpfungs-<br>quotient | Anteile von Verwandtschaftsbeziehungen<br>(Lebenspartner, Eltern, Geschwister) |        |        |        |      |
|----------------------------|--|--------|--------|--------|------|
|                            | bis 25%  | 25-50% | 51-75% | 76-99% | 100% |
| 0- 25%                     | 6,9  | 8,6    | 3,6    | -      | -    |
| 26-50%                     | 10,9   | 30,1   | 23,6   | -      | 11,5 |
| 51-75%                     | 13,9   | 9,7    | 5,5    | 14,3   | 3,1  |
| 76-99%                     | 3,0  | 4,3    | 3,6    | 42,9   | 0,8  |
| 100%                       | 65,3   | 47,3   | 63,6   | 42,9   | 84,7 |
| Anzahl:                    | 101  | 93     | 55     | 7      | 131  |

N=387; Chi-Quadrat=130.02; df=16; Cramers`V=.392; Pearsons`R=.24,  
P < 0.0001

**Tabelle 5.2: Ausschöpfungsquotient und Bekanntheitsdauer (Spaltenprozente)**

| Ausschöpfungs-<br>quotient | Anteile Bekanntheitsdauer "Kennen seit Geburt"* |        |        |        |       |
|----------------------------|---|--------|--------|--------|-------|
|                            | bis 25%   | 25-50% | 51-75% | 76-99% | 100%  |
| 0- 25%                     | 5.0   | 4.4    | 3.0    | -      | -     |
| 26-50%                     | 15.1  | 26.4   | 18.2   | 100.0  | -     |
| 51-75%                     | 8.8   | 5.5    | 15.2   | -      | -     |
| 76-99%                     | 2.5   | 4.4    | 9.1    | -      | -     |
| 100%                       | 68.5  | 59.3   | 54.5   | -      | 100.0 |
| Anzahl:                    | 163   | 54     | 18     | 1      | 24    |

N = 387; Chi-Quadrat=30.64; df=16; Cramers`V=.140; Pearsons`R=.08,  
P < 0.0381

\* = Diese Kategorie trifft nur für Familienmitglieder zu und restringiert diese Analyse als Bestandteil der Anteile familialer Beziehungen.

**Tabelle 5.3: Ausschöpfungsquotient und Kontaktfrequenz**

| Ausschöpfungs-<br>quotient | Anteile von Netzpersonen mit häufiger Kontakt-<br>frequenz (täglich, mehrmals die Woche) |        |        |        |      |
|----------------------------|--|--------|--------|--------|------|
|                            | bis 25%  | 25-50% | 51-75% | 76-99% | 100% |
| 0- 25%                     | 8.8  | 2.7    | 7.3    | 26.7   | 1.9  |
| 26-50%                     | 2.9  | 37.3   | 20.0   | 20.0   | 11.5 |
| 51-75%                     | 8.8  | 4.0    | 18.2   | 6.7    | 6.7  |
| 76-99%                     | 2.9  | 10.7   | 1.8    | 13.3   | 0.5  |
| 100%                       | 76.5   | 45.3   | 52.7   | 33.3   | 79.3 |
| Anzahl:                    | 34   | 75     | 55     | 15     | 208  |

N=387; Chi-Quadrat=94.26; df=16; Cramers `V=.490; Pearsons`R=.18,  
P < 0.0491

**Tabelle 5.4: Ausschöpfungsquotient und emotionale Nähe**

| Ausschöpfungs-<br>quotient | Anteile sehr hoher emotionaler Nähe |        |        |        |      |
|----------------------------|-------------------------------------|--------|--------|--------|------|
|                            | bis 25%                             | 25-50% | 51-75% | 76-99% | 100% |
| 0- 25%                     | 5.1                                 | 10.0   | -      | -      | 1.4  |
| 26-50%                     | 12.7                                | 24.5   | 24.5   | 28.6   | 11.3 |
| 51-75%                     | 7.6                                 | 10.9   | 16.3   | -      | 3.5  |
| 76-99%                     | 6.3                                 | 2.7    | 2.0    | 28.6   | 1.4  |
| 100%                       | 68.4                                | 51.8   | 57.1   | 42.9   | 82.4 |
| Anzahl:                    | 79                                  | 110    | 28     | 7      | 142  |

N=387; Chi-Quadrat=59.58; df=16; Cramers `V=.392, Pearsons`R=.18,  
P < 0.0457

**Tabelle 6: Ausschöpfungsanteile im Verhältnis von genannten Netzadressen und vorliegenden Netzrecords (in %)**

| Ausschöpfungs-<br>quotient | Anzahl genannter Netzadressen |      |      |      |      |      |      |    |     |     |
|----------------------------|-------------------------------|------|------|------|------|------|------|----|-----|-----|
|                            | 1                             | 2    | 3    | 4    | 5    | 6    | 7    | 8  | 9   | 10  |
| bis 25%                    | -                             | -    | -    | 24.1 | 7.4  | 15.4 | 11.1 | -  | -   | 100 |
| 26 -50%                    | -                             | 54.2 | 28.6 | 31.0 | 29.6 | 38.5 | 22.2 | -  | 100 | -   |
| 51 -75%                    | -                             | -    | 46.4 | 31.0 | 22.2 | 23.1 | 44.4 | 75 | -   | -   |
| 76 -99%                    | -                             | -    | -    | -    | 25.9 | 23.1 | 22.2 | 25 | -   | -   |
| 100%                       | 100*                          | 45.8 | 25.0 | 13.8 | 14.8 | -    | -    | -  | -   | -   |
| Anzahl:                    | 61                            | 59   | 56   | 29   | 27   | 13   | 9    | 4  | 1   | 1   |

Chi-Quadrat=134.27, df=24, Cramers`V=.359; Pearsons`R=-.304,  $p < .00001$

\* = Dieser Anteil ist bedingt durch das Auswahlkriterien, daß mindestens eine Netzperson geantwortet hat.

**Tabelle 7: Vergleich der Netzpersonen-Populationen**

| Variable       | Befragungsgruppe       |                     |           |
|----------------|------------------------|---------------------|-----------|
|                | Ausgangspopulation:    | Rücklaufpopulation: | Differenz |
|                | Alle von Ego           | Alle Netzperson     |           |
|                | benannten Netzpersonen | mit Rückantwort     |           |
|                | in %                   | in %                |           |
|                | N = 2730               | N = 557             |           |
| männlich       | 54.9                   | 56.4                | + 1.5     |
| weiblich       | 45.1                   | 43.2                | - 1.9     |
| Volksschule    | 51.9                   | 40.7                | - 11.2    |
| Realschule     | 23.0                   | 25.4                | + 2.4     |
| Abitur         | 11.7                   | 16.6                | + 4.9     |
| Hochschule     | 6.7                    | 12.8                | + 6.1     |
| Ohne Abschluß/ |                        |                     |           |
| Sonstiges      | 5.9                    | 4.5                 | - 1.4     |
| bis 20 Jahre   | 8.0                    | 7.1                 | - 0.9     |
| 21 -30 Jahre   | 19.8                   | 17.9                | - 1.9     |
| 31 -40 Jahre   | 19.2                   | 15.6                | - 3.6     |
| 41 -50 Jahre   | 19.7                   | 17.4                | - 2.3     |
| 51 -60 Jahre   | 15.8                   | 19.7                | + 3.9     |
| 61 und älter   | 17.6                   | 22.4                | + 4.8     |

**Tabelle 8: Validierung von Proxy-Daten der Alt-Neustadt-Studie**

| Soziodemo-<br>graphische<br>Variable | Intra-<br>klassen-<br>Korrelation | Prozentuale<br>Übereinstimmung |
|--------------------------------------|-----------------------------------|--------------------------------|
| Alter                                | .78                               | 31 (73)                        |
| Berufsangabe                         | -                                 | 64                             |
| Religion                             | .76                               | 88                             |
| Parteipräferenz                      | .49                               | 68                             |

Quelle: Pappi/Wolf 1984: S. 288, Spalte 1 und 3. In Klammern: Prozentuale Übereinstimmung bei einer zulässigen Abweichung von plus/minus 2 Jahren. (N=113).

**Tabelle 8.1: Validierung von Proxy-Daten der DFG-Studie Hohenheim**

| Soziodemo-<br>graphische<br>Variable | Test-Retest<br>Koeffizient | Prozentuale<br>Übereinstimmung |
|--------------------------------------|----------------------------|--------------------------------|
| Geschlecht*                          | .96                        | 98                             |
| Alter**                              | .90                        | 78                             |
| Bildung***                           | .48                        | 62                             |

(N = 261) \* = Phi; \*\* = Pearsons`R; \*\*\* = Spearman`s R

**Tabelle 9: Übereinstimmung von Eigenangaben (Test-Retest-Korrelation)**

|                | Prozentuale<br>Übereinstimmung | Cramers` V | Test-Retest-<br>Koeffizient |
|----------------|--------------------------------|------------|-----------------------------|
| Geschlecht*    | 91                             | .82        | .82                         |
| Alter**        | 76                             | .72        | .79                         |
| Bildung***     | 47                             | .38        | .42                         |
| Wahlabsicht*** | 65                             | .45        | .44                         |

Gemessen als: \* = Phi; \*\* = Pearsons`R; \*\*\* = Spearman`s R

**Tabelle 9.1: Korrigierte Test-Retest-Korrelation der Eigenangaben**

|             | Prozentuale<br>Übereinstimmung | Cramers` V | Test-Retest-<br>Koeffizient |
|-------------|--------------------------------|------------|-----------------------------|
| Geschlecht  | 100                            | 1.00       | 1.00                        |
| Alter       | 94                             | .78        | .88                         |
| Bildung     | 64                             | .39        | .54                         |
| Wahlabsicht | 67                             | .46        | .46                         |

Zur Verwendung der Maße vgl. Tabelle 8

**Tabelle 10.1: Verteilung der Kontextmerkmale (in % aller Befragten; N=392)**

| Merkmal                               | In % | Schiefe |
|---------------------------------------|------|---------|
| Formale Art der sozialen Beziehung    |      | .304    |
| LebenspartnerIn                       | 23.2 |         |
| Familiale Relationen                  | 32.0 |         |
| Freunde                               | 23.7 |         |
| Bekannte                              | 21.1 |         |
| Bekanntheitsdauer                     |      | .901    |
| bis zu 3 Jahren                       | 10.0 |         |
| 4 bis 10 Jahre                        | 23.1 |         |
| 11 bis 20 Jahre                       | 14.7 |         |
| 21 Jahre und länger                   | 52.2 |         |
| Kontakthäufigkeit und Kontaktfrequenz |      | .748    |
| täglich, mehrmals die Woche           | 66.8 |         |
| seltener                              | 33.2 |         |
| Affektive bzw. emotionale Nähe        |      | 1.046   |
| hoch                                  | 57.0 |         |
| mittel                                | 21.4 |         |
| niedrig                               | 21.6 |         |



Tabelle 10.2: Kontexteffekte auf die Güte retrospektiver Daten

| Kontextmerkmal    | Chi-Quadrat | Freiheits-<br>grade | Cramer's R | Korrelations-<br>koeffizient | p-Wert |
|-------------------|-------------|---------------------|------------|------------------------------|--------|
| Alter (N=388)     |             |                     |            |                              |        |
| Beziehungsart     | 6.90        | 3                   | .133       | .117                         | .021   |
| Bekanntheitsdauer | 4.06        | 3                   | .101       | -.027                        | .059   |
| Kontakthäufigkeit | 0.09        | 1                   | .002       | .002                         | .978   |
| Affektive Nähe    | 2.44        | 2                   | .087       | .036                         | .355   |
| Bildung (N=338)   |             |                     |            |                              |        |
| Beziehungsart     | 3.43        | 3                   | .100       | .097                         | .071   |
| Bekanntheitsdauer | 2.47        | 3                   | .085       | -.084                        | .125   |
| Kontakthäufigkeit | 0.69        | 1                   | .044       | .045                         | .416   |
| Affektive Nähe    | 2.32        | 2                   | .082       | .078                         | .134   |

Tabelle 11: Paarvergleich von Reliabilität und Proxy-Angabe

| Kombination<br>der<br>Variablenpaare            | Verteilung für die ..... |             |                  |      |
|---|--------------------------|-------------|------------------|------|
|   | Variable Alter           |             | Variable Bildung |      |
|   | abs.<br>(N=392)          | in %        | abs.<br>(N=353)  | in % |
| Reliabilität und<br>Proxy kongruent             | 100 (271)                | 25.6 (69.7) | 185              | 54.6 |
| Reliabilität kongruent und<br>Proxy inkongruent | 207 (70)                 | 53.1 (18.0) | 79               | 23.3 |
| Reliabilität inkongruent<br>und Proxy kongruent | 20 (36)                  | 5.1 (9.3)   | 42               | 12.4 |
| Reliabilität und Proxy<br>inkongruent           | 63 (12)                  | 16.2 (3.1)  | 33               | 9.7  |

In Klammern: Angaben für eine Abweichung von plus/minus einem Jahr

**Tabelle 12: Kreuztabelle der kombinatorischen Paarvergleiche**

| Variable Alter  | Variable Bildung |      |     |     |
|-----------------|------------------|------|-----|-----|
|                 | E=1              | E=1  | E=0 | E=0 |
|                 | N=1              | N=0  | N=1 | N=0 |
| <b>E=1, N=1</b> |                  |      |     |     |
| abs.            | 58               | 22   | 5   | 2   |
| in %            | 17.1             | 6.5  | 1.5 | 0.6 |
| <b>E=1, N=0</b> |                  |      |     |     |
| abs.            | 101              | 46   | 17  | 15  |
| in %            | 29.8             | 13.6 | 5.0 | 4.4 |
| <b>E=0, N=1</b> |                  |      |     |     |
| abs.            | ..6              | 4    | 5   | 3   |
| in %            | 1.8              | 1.2  | 1.5 | 0.9 |
| <b>E=0, N=0</b> |                  |      |     |     |
| abs.            | 20               | 7    | 15  | 13  |
| in %            | 5.9              | 2.1  | 4.4 | 3.8 |

E steht für die Eigenangaben, N steht für die Netzpersonen-Proxys. Der Wert Eins bedeutet Kongruenz, der Wert Null bedeutet Inkongruenz (Fallzahl=339 Relationen). (Chi-Quadrat=145.67, df=9; Cramers`V=.211; Spearman-Koeffizient=.2758,  $p < .0001$ )

**Tabelle 13: Korrelationsmaße zur Güte von soziodemographischen Proxy-Daten bei einer zulässigen Altersabweichung von plus/minus einem Jahr**

| Variable | Eigenangaben | Proxy-Angaben | p-Wert   | N   |
|----------|--------------|---------------|----------|-----|
| Bildung* | .74          | .69           | < .00001 | 353 |
| Alter    | .91 #        | .94           | < .00001 | 390 |

\* betrachtet als ordinalskalierte Variable, mit Spearman-Rangkorrelation berechnet

**Tabelle 14:** Kreuztabelle der kombinatorischen Paarvergleiche mit  
**modifiziertem Altersabgleich**  
 (in % aller Befragten, N=347)

|                       | Variable Bildung |            |            |            |
|-----------------------|------------------|------------|------------|------------|
|                       | E=1<br>N=1       | E=1<br>N=0 | E=0<br>N=1 | E=0<br>N=0 |
| <b>Variable Alter</b> |                  |            |            |            |
| E=1, N=1              |                  |            |            |            |
| abs.                  | 150              | 56         | 17         | 18         |
| in %                  | 43.2             | 16.1       | 4.9        | 5.2        |
| E=1, N=0              |                  |            |            |            |
| abs.                  | 27               | 20         | 7          | 9          |
| in %                  | 7.9              | 5.8        | 2.0        | 2.6        |
| E=0, N=1              |                  |            |            |            |
| abs.                  | 13               | 1          | 15         | 3          |
| in %                  | 3.7              | 0.3        | 4.3        | 0.9        |
| E=0, N=0              |                  |            |            |            |
| abs.                  | 3                | 2          | 3          | 3          |
| in %                  | 0.9              | 0.6        | 0.9        | 0.9        |

E steht für die Eigenangaben, N steht für die Netzpersonen-Proxys. Der Wert 1 bedeutet Kongruenz, der Wert 0 bedeutet Inkongruenz (Fallzahl=339 Relationen).

(Chi-Quadrat= 59.80, df=9; Cramers`V=.239; Spearman-Koeffizient=.246,  $p < .0001$ )